

Pénzügyi piacok fundamentális és technikai elemzése

PhD Értekezés tézisei

Torma Balázs

Témavezető: Gerencsér László, DSc



Eötvös Loránd Tudományegyetem,
Informatikai Kar
(ELTE IK)

Informatikai Doktori Iskola,
Az informatika alapjai és módszertana doktori program,
Elnök: Prof. Demetrovics János, MTA rendes tagja

Magyar Tudományos Akadémia,
Számítástechnikai és Automatizálási Kutatóintézet
(MTA SZTAKI)

Budapest, 2010

Bevezető

Hazai és globális gazdaságok stabilitásának biztosítása érdekében elengedhetetlen, hogy értsük a pénzügyi piacok működési mechanizmusait. A piaci résztvevők motivációi, tőkekitettségek nagyon különböző, illetve gyakran különböző kereskedési szabályok vonatkoznak rájuk. Ezt a magas fokú heterogenitást multiágens piacmodellek (MÁP modellek) segítségével írhatjuk le, mivel ezek a modellek az ágensok viselkedését és a gazdasági folyamatokat részleteiben modellezzik.

A MÁP-modellek egy fontos alkalmazása a nem megfigyelhető gazdasági folyamatok vizsgálata megfigyelt adatok alapján. A gazdasági változókra való statisztikai következtetést nagy mértékben nehezíti e modellek komplexitása. A nehézséget elkerülendő, matematikailag kezelhető ágensmodellek fejlesztése indult meg a közelmúltban, ld. Boswijk et al. (2007); Kozhan and Salmon (2009). A kezelhetőségért viszont magas árat kell fizetni: ezek a modellek nagyon absztraktak, változóikat magas fokú aggregátság jellemzi illetve irreális feltételeket írnak elő - azaz a piac fontos jellemzőit hanyagolják el. Ennek az a következménye, hogy gazdaságilag csak korlátozottan értelmezhetőek azok az eredmények, amelyeket ezen egyszerű modellek adatokra való illesztésével nyerünk, ld. Amilon (2008).

Ebben az értekezésben egy részletes MÁP-modellt javasolok és matematikailag kezelhető technikai (ökonometriai) modellek alapján fejlesztett statisztikai algoritmusokat mutatok be, illetve technikai modellek segítségével elemzem a MÁP-modellben szereplő egyes kulcsfontosságú gazdasági tényezőket. A statisztikai elemzés fő matematikai eszköze a modellidentifikáció Legnagyobb Valószínűség Elve alapján, ld. Ljung and Söderström (1983); Benveniste et al. (1990); Spall (2003). Bemutatok egy új, hatékony off-line identifikációt megvalósító lokális optimalizálási algoritmust, illetve néhány fontos technikai modellen alapuló valós idejű, rekurzív identifikáló és változás-detektáló algoritmust.

Tudományos eredmények

1. tézis: Változás-detektálás fundamentális modellekben technikai modellekkel

A MÁP modellekben általában két fő ágenstípust különböztetünk meg. A technikai kereskedők piaci várakozásait a jelenlegi trend alapján határozzák meg, míg a fundamentalisták hiedelme a jövőbeli áráról a vállalatról érkező hírek alapján alakul. A piaci struktúrát a piaci összvagynak a két ágenstípus közötti megoszlásával definiálhatjuk. Kirman (1995), Brock and Hommes (1998) and Lux (1998) úttörő műveiben a piacstruktúra időbeli fejlődését az egyes ágenstípusokhoz tartozó stratégiák múltbeli nyereségessége határozza meg. Tehát ezek a modellek azt sugallják, hogy a jövőre vonatkozó elvárások csak a múltban realizált profittól függenek. Ezzel szemben a javasolt MÁP-modellben megengedjük, hogy az elvárások külső gazdasági hatásoktól is függhessenek és így a javasolt MÁP-modellben a piacstruktúra (exogén) paraméter.

Mivel piaci hírek megjelenésekor a pénzügyi elemzők még több hírt generálnak, a hírek érkezési folyamatát pozitív visszacsatolású (öngerjesztő) folyamatokkal modellezzük.

- *Javasolok egy fundamentális részvénypiac-modellt, amely az irodalomban fellelhető MÁP-modellkoncepciókat egészíti ki információ érkezési folyamatokkal, amelyeket egy ún. diszkrét Hawkes folyamattal modellezek, ld. Hawkes (1971). Numerikus kísérletekben megmutatom, hogy a modell reprodukálja a piaci idősorokban fellelhető tipikus mintákat, az ún. stilizált tényeket a fundamentalisták és technikai kereskedők fix aránya mellett is, szemben klasszikus modellekkel.*

A tőzsdemodell validálásának egy további lépéseként a generált árfolyamatot egy széles körben elfogadott ökonometriai modell, a General Autoregressive Heteroscedastisity (GARCH) modell (ld. Engle (1982); Bollerslev (1986)) segítségével vizsgálom:

$$\begin{aligned} r_t &= \sqrt{h_t} \varepsilon_t, \\ h_t &= \sigma_0 + \varphi r_{t-1}^2 + \psi h_{t-1}, \end{aligned}$$

ahol r_t a log-hozamokat jelöli, $\varepsilon_t \sim \mathcal{N}(0, 1)$ függetlenek, $\vartheta = (\sigma_0, \varphi, \psi)^T$ a valós paramétervektor.

Egy monoton összefüggést létesítek a piac struktúrája és a legjobban illeszkedő GARCH modell között:

- *Numerikus kísérletekben megmutatom, hogy $\hat{\phi}$ nő, ha a technikai kereskedők aránya nő és $\hat{\psi}$ nő, ha a fundamentalisták aránya nő, ahol $\hat{\phi}, \hat{\psi}$ a MÁP-modell által generált idősorokra illesztett $GARCH(1,1)$ paramétereket jelöli.*

A fenti kapcsolat által motiválva a $GARCH(1,1)$ modellt használjuk a piacstruktúrára vonatkozó statisztikai következtetésekre. A javasolt fundamentális modell statisztikai elemzésének korlátot szab a modell komplexitása. Ugyanakkor ez az akadály elhárítható ha csupán a modellben (dinamikában) történő változást akarjuk detektálni. Ismert ui., hogy a változás-detektáló algoritmusok tipikusan robusztusak, azaz alkalmazhatók rosszul specifikált modell felhasználásával is, ld. Basseville and Nikiforov (1993).

- *Egy új, GARCH modell szerinti valós idejű változás-detektáló algoritmust mutatok be Gerencsér and Baikovicus (1991, 1992) Minimális Leíróhossz alapú módszereit követve. A változás-detektáló algoritmus tartalmaz egy új, hatékony rekurzív GARCH becslési elemet. Az algoritmust sikeresen alkalmaztam a fenti tőzsdemodellben létrejövő piaci struktúra megváltozásának detektálására. Az algoritmust futtattam valódi árfolyam adatokon is.*

A fenti, a GARCH és MÁP-modell közötti kapcsolat alapján mintegy közgazdaságilag értelmezem a GARCH szerinti változás-jelzéseket:

- *Annak feltételeit tárgyalom, hogy a GARCH alapú változás-detektáló algoritmus a piacstruktúrában bekövetkező változást detektálja.*

A tézishez kapcsolódó kéziratot hamarosan publikáljuk. A rekurzív GARCH becslés algoritmusát bemutató kéziratot publikálásra elfogadták, ld. Gerencsér et al. (2010).

2. tézis: Hatékony off-line modellidentifikáció

A tőzsdemodell által generált idősorokra egy standard kvázi-loglikelihood módszerrel illesztünk egy GARCH modellt, először off-line módon. A felmerülő nemlineáris minimalizálási probléma közismerten rosszul kondicionált.

Az általános optimalizálási probléma a következőképpen fogalmazható meg:

$$\min_{x \in S} f(x),$$

ahol $f : S \rightarrow \mathbb{R}$ egy kétszer folytonosan differenciálható függvény, melynek x^* optimális pontja S belsejében van, ahol $S \subset \mathbb{R}^n$ megengedhető tartományt korlátaival adjuk meg.

- *Egy új, általános nemlineáris nemkonvex optimalizáló algoritmust javasolok, amely vágósík alapú technikák és (kvázi-) Newton módszerek keveréke. Az algoritmust konvex függvényeken, legkisebb négyzetek problémákon és GARCH paraméterbecslésen tesztelem. Numerikus kísérletekben megmutatom, hogy a módszer az említett problémák esetén konvergál és tipikusan sokkal gyorsabb.*

A módszer vágósíkok használatával kerül gyorsan az optimum közelébe és onnét (kvázi) Newton eljárások sebességével tart az optimumhoz. A numerikus tesztek eredményei azt mutatják, hogy a módszer nem érzékeny az x_0 kezdeti értékre, S méretére és zajos célfüggvény esetén annak számos lokális szélsőértékére.

A tézishez kapcsolódó tudományos eredményeket a Torma and G.-Tóth (2010) cikkben publikáltuk.

3. tézis: Hírfolyamatok modellezése

A tőzsdei elemzők körében megfigyelhető nyájeffektus öngerjesztő hatású, ld. pl. Welch (2000): egy cégről megjelenő hírek és jelentések az elemzőket arra ösztökélik, hogy erre a cégre fókuszáljanak, és így növelik a cég elemzők általi lefedettségét. Ez viszont még több hírt eredményez az adott cégről. Ezt a heurisztikát öngerjesztő pontfolyamatokkal modellezzük.

- *Az hírfolyamat dinamikájának modellezésére olyan Hawkes folyamatokat javasolok, melyekben a visszacsatolást az eseményfolyamat és az intenzitás között egy véges dimenziós lineáris rendszerrel írjuk le. Algoritmusokat javasolok a Hawkes folyamatok szimulációjára és hatékony rekurzív identifikációjára. A becslés konvergenciáját numerikus kísérletekben igazolom.*

A javasolt Hawkes folyamat egy olyan $N(t)$ pontfolyamat, melynek $\lambda(t)$ intenzitása kielégíti a következő lineáris állapotegyenletet:

$$\begin{aligned} dx(t) &= -Ax(t)dt + bdN(t), \\ \lambda(t) &= c^T x(t-) + m, \end{aligned}$$

ahol A mátrix, b, c vektorok és m skalár, melyek kielégítik a következő stabilitási feltételt:

$$c^T A^{-1} b < 1.$$

Ebben a modellben az x állapotvektor az elemzők aktivitását reprezentálja.

Vizsgáltam a becslési probléma Fischer információs mátrixát a standard Hawkes modell esetén, amelyben $a = A$, $\sigma = bc$, m skalár paraméterek. Ebben a modellben a stabilitási kritérium $\sigma < a$ -ra egyszerűsödik.

- *Az a és σ paraméterek szerinti aszimptotikus Fischer információ végtelenbe tart, az m paraméter szerinti aszimptotikus Fischer információ nullához tart, ha a paraméterekkel a stabilitási tartomány széléhez tartunk, azaz $a - \sigma \rightarrow 0$. A megfelelően átskálázott intenzitás eloszlásban konvergál a $\Gamma_{\frac{2m}{\sigma_0}, \frac{2}{\sigma_0^2}}$ eloszláshoz, ha a és σ σ_0 -hoz tartanak úgy, hogy $\sigma < a$.*

A tézishez kapcsolódó tudományos eredményeket a Gerencsér et al. (2008b); Prokaj and Torma (2010) cikkekben publikáltuk.

4. tézis: Nagyfrekvenciás tőzsdeadatok regressziós analízise

Értékpapírárak, mint pl. részvényárak egy általános tulajdonsága, hogy az ár egy ún. minimális értéknövekmény vagy tick-méret egész számú többszöröse. A NASDAQ tőzsdén például ez a minimális értéknövekmény $h = 0.01\$$. A minimális értéknövekmény a tőzsdén alkalmazott árazó algoritmusok miatt szükséges, mivel ezek a keresletet és kínálatot egymástól h távolságra levő árszinteken aggregálják. A piaci árat tehát felfoghatjuk aggregálás után megfigyelt árként. Az aggregálásnak ezt a formáját kvantálásnak is nevezzük. A kvantálás miatt akkor veszünk sok információt, ha h nagy az ár volatilitásához képest. Ez teljesül pl. magasfrekvenciás adatok vizsgálatánál, azaz amikor az árat nagy frekvencián

mintavételezzük, mondjuk 100-szor egy másodperc alatt. Napjainkban fontos kutatási terület az ajánlati könyvről érkező adatok vizsgálata, mint pl. az ár dinamikája és a különböző árszinteken megjelenő ajánlati mennyiségek közötti összefüggések keresése.

A fenti probléma által motiválva azt a regressziós problémát vizsgáltam, amelyben a $(\psi) = \psi_n \in \mathbb{R}^d$ véges értékű regresszorfolymat $\vartheta^* \in \mathbb{R}^d$ együtthatóvektorát szeretnénk becsülni kvantált, additív Gauss zajjal terhelt mérésekből. A megfigyelt értékek alakja tehát:

$$y_n = q(\psi_n^T \vartheta^* + e_n),$$

ahol e_n egy független azonos eloszlású 0 várható értékű és $\sigma^2 = (\sigma^*)^2$ szórásnégyszetű Gauss sorozat. A q skalár kvantálót egy olyan leképezésként definiálhatjuk, amely \mathbb{R} -ből egy $\mathcal{Y} \subset \mathbb{R}$ diszkrét, véges vagy megszámlálható halmazba képez, ahol \mathcal{Y} az ún. kvantálási szinteket reprezentálja. A kvantáló $x \in \mathbb{R}$ -hez rendeli annak kvantált változatát: $y = q(x)$.

- *Egy új, hatékony rekurzív algoritmust javasolok a lineáris regresszió együtthatóinak kvantált megfigyeléseiből való becslésére, véges értékészletű regresszor mellett. A módszer konvergenciáját numerikus kísérletekben igazolom.*

Az algoritmusban Expectation Maximization és Markov Chain Monte Carlo technikákat alkalmazunk.

A tézishez kapcsolódó tudományos eredményeket a Gerencsér et al. (2008a) cikkben publikáltuk.

Publikációk

A tézisekhez kapcsolódó publikációk

Folyóíratcikkek

- Gerencsér, L., Kmecs, I., Torma, B., 2008a. Quantization with adaptation, estimation of gaussian linear models. Communications in Information and Systems 8 (The Brockett Legacy Special Issue. In Honor of Roger W. Brockett in honor of his 75th birthday (guest eds.: John Baillieul, John S. Baras, Anthony Bloch, P.S. Krishnaprasad and Jan C. Willems)), 223–244.
- Prokaj, V., Torma, B., 2010. Identification of almost unstable Hawkes processes. Publicationes Mathematicae Debrecen, forthcoming. Impact factor: 0.137
- Torma, B., G.-Tóth, B., 2010. An efficient descent direction method with cutting planes. Central European Journal of Operations Research, forthcoming. (Impact factor first published later in 2010)

Konferenciatickek

- Gerencsér, L., Matias, C., Vágó, Z., Torma, B., Weiss, B., 2008b. Self-exciting point processes with applications in finance and medicine. In proceedings of The 18th International Symposium on Mathematical Theory of Networks and Systems, Blacksburg, Virginia, USA.
- Gerencsér, L., Orlovits, Z., Torma, B., 2010. Recursive estimation of GARCH processes. The 19th International Symposium on Mathematical Theory of Networks and Systems, Budapest, Hungary, forthcoming.

Egyéb cikkek

- Kurucz, M., Benczúr, A., Kiss, T., Nagy, I., Szabó, A., Torma, B., 2007. KDD Cup 2007 Task 1 Winner Report, SIG-KDD Explorations 9(2), 53–56.
- Gerencsér, L., Manno, S., Torma, B. 2010. Reconstruction of sound from a phonograph cylinder using non-contact photographic metrology, Journal of the Audio Engineering Society, due to submit.

Hivatkozások

- Amilon, H., 2008. Estimation of an adaptive stock market model with heterogeneous agents. *Journal of Empirical Finance* 15 (2), 342 – 362.
- Basseville, M., Nikiforov, I., 1993. *Detection of Abrupt Changes: Theory and Application*. Prentice Hall.
- Benveniste, A., Métivier, M., Priouret, P., 1990. *Adaptive algorithms and stochastic approximations*. Springer-Verlag, Berlin.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31 (3), 307–327.
- Boswijk, P., Hommes, C., Manzan, S., 2007. Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control* 31, 1938–1970.
- Brock, W. A., Hommes, C. H., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control* 22, 1235–1274.
- Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50 (4), 987–1007.
- Gerencsér, L., Baikovicius, J., 1991. Change-point detection using stochastic complexity. identification and system parameter estimation. *Selected papers from the 9th IFAC-IFORS Symposium on Budapest* 1, 73–78.
- Gerencsér, L., Baikovicius, J., 1992. Change-point detection as model selection. *Informatica* 3, 3–20.
- Hawkes, A., 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58, 83–90.
- Kirman, A., 1995. The behaviour of the foreign exchange market. *Bank of England Quarterly Bulletin* 15, 286–293.
- Kozhan, R., Salmon, M., 2009. Uncertainty aversion in a heterogeneous agent model of foreign exchange rate formation. *Journal of Economic Dynamics and Control* 33 (5), 1106 – 1122.
- Ljung, L., Söderström, T., 1983. *Theory and practice of recursive identification*. The MIT Press.
- Lux, T., 1998. The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distribution. *Journal of Economic Behavior and Organization* 33, 143–165.
- Spall, J. C., 2003. *Introduction to Stochastic Search and Optimization*. John Wiley & Sons, Inc., New York, NY, USA.
- Welch, I., 2000. Herding among security analysts. *Journal of Financial Economics* 58, 369–396.